3 Lecture 3

3.1 Sum, direct sum, quotient

Theorem 3.1. $\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$

Proof. Let $u_1, \dots, u_k$ be a basis of $V_1 \cap V_2$. By Corollary 2.16, we can extend $u_1, \dots, u_k$ to $u_1, \dots, u_k, v_1, \dots, v_j$ a basis of V_1 and to $u_1, \dots, u_k, w_1, \dots, w_l$ to be a basis of V_2. We claim that $B = \{u_1, \dots, u_k, v_1, \dots, v_j, w_1, \dots, w_l\}$ is a basis of $V_1 + V_2$.

Let $x \in V_1$ and $y \in V_2$. Then $x = \lambda_1 u_1 + \cdots + \lambda_k u_k + b_1 v_1 + \cdots + b_j v_j$ and $y = \mu_1 u_1 + \cdots + \mu_k u_k + c_1 w_1 + \cdots + c_l w_l$ thus $x + y = (\lambda_1 + \mu_1) u_1 + \cdots + (\lambda_k + \mu_k) u_k + b_1 v_1 + \cdots + b_j v_j + c_1 w_1 + \cdots + c_l w_l$. Hence we see that $u_1, \dots, u_k, v_1, \dots, v_j, w_1, \dots, w_l$ span $V_1 + V_2$.

To show that B is linearly independent, we suppose $a_1 u_1 + \cdots + a_k u_k + b_1 v_1 + \cdots + b_j v_j + c_1 w_1 + \cdots + c_l w_l = 0$. Then $b_1 v_1 + \cdots + b_j v_j = -a_1 u_1 - \cdots - a_k u_k - c_1 w_1 - \cdots - c_l v_l \in V_2$. Since $b_1 v_1 + \cdots + b_j v_j \in V_1$, we have $b_1 v_1 + \cdots + b_j v_j \in V_1 \cap V_2$. Thus there exists $\lambda_1, \dots, \lambda_k$ such that $\lambda_1 u_1 + \cdots + \lambda_k u_k + b_1 v_1 + \cdots + b_j v_j = 0$. Since $u_1, \dots, u_k, v_1, \dots, v_j$ is a basis of V_1 we have $b_1 = \cdots = b_j = 0$. Similarly we can show that $c_1 = \cdots = c_l = 0$. Thus we have $a_1 u_1 + \cdots + a_k u_k = 0$. Since $u_1, \dots, u_k$ is a basis of $V_1 \cap V_2$, we have $a_1 = \cdots = a_k = 0$.

In particular, $\dim(V_1 + V_2) = k + j + l = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2)$. □

Remark 3.2. In this proof, we start with a basis of $V_1 \cap V_2$ and extend it to be a basis of V_1 and V_2 respectively. One <u>cannot</u> do it the other way around by directly taking the intersection of bases of V_1 and V_2. The intersection is not necessarily a basis of $V_1 \cap V_2$ as can be seen in the following example. Let $V_1 = \operatorname{span}(v_1, v_2)$ and $V_2 = \operatorname{span}(v_3, v_4)$. We have $\{v_1, v_2\} \cap \{v_3, v_4\} = \emptyset$. However, $V_1 \cap V_2 = \operatorname{span}(v_3)$.

Definition 3.3. We say that $V_1 + \cdots + V_m$ is a **direct sum** if any $x \in V_1 + \cdots + V_m$ can be written as $v_1 + \cdots + v_m$ *uniquely* for $v_i \in V_i$, $1 \le i \le m$. In this case, we write $V_1 \oplus \cdots \oplus V_m$ for $V_1 + \cdots + V_m$.

We want to introduce the summation notation here which will be used in this notes.

Notation 3.4. If $x_1, \dots, x_n$ are things that can be added together e.g. numbers, vectors, functions, matrices etc, then we write

$$\sum_{i=1}^{n} x_i$$

for

$$x_1 + \cdots + x_n.$$

The symbol reads "summation of x_i for i from 1 to n".

The following Proposition shows that a direct sum is a sum of subspaces which has "no overlap".

Proposition 3.5. *Let V be a finite dimensional vector space. The following are equivalent:*

(i) The only way to write $0 = v_1 + \cdots + v_m$ where $v_i \in V_i$ is by taking $v_i = 0$ for all $1 \leq i \leq m$.

(ii) $V_1 + \cdots + V_m$ is a direct sum.

(iii) $V_i \cap \sum_{j:j\neq i} V_j = \{0\}$ for all $1 \leq i \leq m$.

(iv) $\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m$.

Proof. (i) $\implies$ (ii) Suppose $x = \sum_{i=1}^m v_j = \sum_{j=1}^m w_j$ where $v_j, w_j \in V_i$ for any j. Then $0 = \sum_{j=1}^m (v_j - w_j)$. By (i) we have $v_j = w_j$ for all $1 \leq j \leq m$ and hence the decomposition is unique.

(ii) $\implies$ (iii) Let $v \in V_i \cap \sum_{j:j\neq i} V_j$. Then $v \in V_i$ and $v = \sum_{j:j\neq i} v_j$ for $v_j \in V_j$, $j \neq i$. Then $0 = v_1 + \cdots + v_{i-1} - v + v_{j+1} + \cdots + v_m$. By (ii), we have $v = 0$.

(iii) $\implies$ (iv) For any $1 \leq j \leq m$, let $\{w_i^{(j)}\}_{1\leq i\leq k_j}$ be a basis of V_j. We claim that $\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$ is a basis of $V_1 + \cdots + V_m$. Suppose $\sum_{j=1}^m \sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)} = 0$ for $a_i^{(j)} \in F$. Then for $1 \leq j \leq m$, $\sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)} = -\sum_{l\neq j} \sum_{i=1}^{k_l} a_i^{(j)} w_i^{(l)} \in \sum_{l:l\neq j} V_l$. Since $\sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)} \in V_j$, we have $\sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)} \in V_j \cap \sum_{l:l\neq j} V_l = \{0\}$. Thus $\sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)} = 0$. Since $\{w_i^{(j)}\}_{1\leq i\leq k_j}$ is a basis of V_j, we have $a_i^{(j)} = 0$ for all $1 \leq i \leq k_j$. Since this is true for all $1 \leq j \leq m$, we have $\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$ is linearly independent. Since any $v \in V_1 + \cdots + V_m$ can be written as $v = v_1 + \cdots + v_m$ and $v_j \in V_j$, we have $v_j = \sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)}$. Hence $v = \sum_{j=1}^m \sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)}$. Therefore, $V_1 + \cdots + V_m = \mathrm{span}\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$.

(iv) $\implies$ (i) For any $1 \leq j \leq m$, let $\{w_i^{(j)}\}_{1\leq i\leq k_j}$ be a basis of V_j. As in previous step we can show that $V_1 + \cdots + V_m = \mathrm{span}\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$. Since $\dim(V_1 + \cdots + V_m) = \dim V_1 + \cdots + \dim V_m = k_1 + \cdots + k_m$, we must have $\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$ is linearly independent since otherwise the spanning set $\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$ strictly contains a basis (by linear dependence and Lemma 2.11) which has elements $< \dim(V_1 + \cdots + V_m)$ a contradiction. Now suppose $0 = v_1 + \cdots + v_m$ where $w_i \in V_i$ then we can write $v_j = \sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)}$ and $0 = \sum_{j=1}^m \sum_{i=1}^{k_j} a_i^{(j)} w_i^{(j)}$. Since $\{w_i^{(j)}\}_{1\leq i\leq k_j, 1\leq j\leq m}$ is linearly independent, we have $v_j = 0$ for all $1 \leq j \leq m$. $\square$

Corollary 3.6. *$V_1 + V_2$ is a direct sum if and only if $V_1 \cap V_2 = \{0\}$.*

Example 3.7. $V_1 = \mathrm{span}(v_1, v_2) = \{(a_1, a_2, 0)^t : a_1, a_2 \in \mathbb{R}\}$

$V_2 = \mathrm{span}(v_3, v_4) = \{(a_3, a_3, a_4)^t : a_3, a_4 \in \mathbb{R}\}$.

$V_1 + V_2 = \mathbb{R}^3$ is not a direct sum: $0 = (v_1 + v_2) - v_3$ and $V_1 \cap V_2 = \mathrm{span}(v_3)$ and $\dim(V_1 + V_2) = 3 < 4 = \dim V_1 + \dim V_2$.

$V_1 + W_1 = \mathbb{R}^3$ is a direct sum.

Lemma 3.8. *Let V be a finite dimensional space and $W \subset V$ be a subspace. Then there exists a subspace U of V such that $V = W \oplus U$.*

Proof. Let $w_1, \ldots, w_m$ be a basis of W. By Corollary 2.16, it can be extended to $w_1, \ldots, w_m, v_1, \ldots, v_{n-m}$ a basis of V. Let $U = \operatorname{span}(v_1, \ldots, v_{n-m})$. Then we have $V = W + U$. Since $\dim W + \dim U = \dim V$, we have that $V = W \oplus U$. $\square$

Definition 3.9. U is called the **complement** of W in V.

3.2 Quotient space

Let $W \subset V$ be a subspace. We are going to define a vector space V/W which is obtained from V by "collapsing" planes parallel to W to a point.

Definition 3.10. Let V be a vector space over F and $W \subset V$ be a subspace. We define the **coset** (**translate**) of W containing $x \in V$ to be $x + W = \{x + w : w \in W\}$. The **quotient space** V/W is the set of all cosets of W that is $V/W = \{x + W : x \in V\}$.

The **canonical projection** associated to a quotient space is $\pi : V \to V/W$, $\pi(x) = x + W$.

Lemma 3.11. *$x + W = y + W$ if and only if $x - y \in W$.*

Proof. Since $x + W = y + W$, we have $x \in y + W$. Then by definition of $y + W$, we have $x - y \in W$.

If $x - y \in W$, then there is $w \in W$ such that $x = y + w$. Hence for any $x_1 \in x + W$, $x_1 = x + w_1 = y + (w + w_1) \in y + W$. Similarly, for any $y_1 \in y + W$, $y_1 = y + w_2 = x - w + w_2 \in x + W$. $\square$

Definition 3.12. We define the addition and scalar multiplication on V/W as $(x + W) + (y + W) := (x + y) + W$ and $c(x + W) = (cx) + W$ for $x, y \in V, c \in F$.

It is not clear whether the definition above are well-defined, since it defines the addition of two cosets $(x + W)$ and $(y + W)$ as the coset $(x + y) + W$ which depends on the choice of x and y in $x + W$ and $y + W$. If we pick different representatives, does the addition give the same coset? This is answered affirmatively in the following lemma.

Lemma 3.13. *Addition and scalar multiplication on V/W are well-defined.*

Proof. For any $x_1 \in x + W$, $y_1 \in y + W$, we have $x_1 - x \in W$ and $y_1 - y \in W$. Thus $(x_1 + y_1) - (x + y) = (x_1 - x) + (y_1 - y) \in W$ and $cx_1 - cx = c(x_1 - x) \in W$. Thus we have $(x_1 + y_1) + W = x + y + W$, $(cx_1) + W = (cx) + W$. $\square$

Example 3.14. $W = \operatorname{span}((0, 1)^t) = \{(0, a)^t : a \in F\} \subset F^2$.

For any $x \in F^2$, $x + W$ the vertical line through x.

$V/W =$ the set of all vertical lines in F^2.

$\pi(x) = x + W$ the vertical line through x.

For example

$0 + W = W$

$x = (1, 0)^t \in F^2$ and $y = (2, 1)^t$

$x + W = \{(1, a)^t : a \in F\}$

$y + W = \{(2, b+1)^t : b \in F\} = \{(2, a)^t : a \in F\} = (2, 0)^t + W$
$(x + y) + W = \{(3, a+1)^t : a \in F\} = \{(3, b)^t : b \in F\}$
$(-x) + W = \{(-1, a)^t : a \in F\}$
In general $(x_1, x_2)^t + W = \{(x_1, a)^t : a \in F\}$ and we have
$((x_1, x_2)^t + W) + ((y_1, y_2)^t + W) = \{(x_1 + x_2, a)^t : a \in F\}$
$c((x_1, x_2)^t + W) = \{(cx_1, a)^t : a \in F\}$.

In other words, adding two cosets in this particular case is just adding the first component and scalar multiplying a coset is just scalar multiplying the first component.

Let $U = \operatorname{span}((1, 0)^t)$. The addition and scalar multiplication are also purely on the first component. Thus the vector space structure of U is "the same as" that of V/W or in math notation $U \cong V/W$.

Proposition 3.15. *V/W is a vector space over F.*

Proof. We check the axioms of a vector space. For $x, y, z \in V$, $a, b \in F$ we have
I: $(x + W) + (z + W) = (x + z) + W = (z + x) + W = (z + W) + (x + W)$.
II: $((x + W) + (y + W)) + (z + W) = ((x + y) + W) + (z + W) = ((x + y) + z) + W = (x + (y + z)) + W = (x + W) + ((y + W) + (z + W))$.
III: $(x + W) + (0 + W) = (x + 0) + W = x + W$.
IV: $(x + W) + ((-x) + W) = (x + (-x)) + W = 0 + W$.
V: $(a + b)(x + W) = ((a + b)x) + W = (ax + bx) + W) = (ax) + W) + (bx) + W = a(x + W) + b(x + W)$.
VI: $(ab)(x + W) = (ab)x + W = a(bx) + W = a(bx + W) = a(b(x + W))$.
VII: $a((x + W) + (y + W)) = a((x + y) + W)) = a(x + y) + W = (ax + ay) + W = (ax + W) + (ay + W) = a((x + W)) + a((y + W))$.
VIII: $1(x + W) = (1x) + W = x + W$. $\square$

Theorem 3.16. *If V is a finite dimensional vector space and W is a subspace of V, then*

$$\dim V/W = \dim V - \dim W.$$

Proof. Let $w_1, \ldots, w_m$ be a basis of W. By Corollary 2.16, we can extend it to $w_1, \ldots, w_m, v_1, \ldots, v_{n-m}$ a basis of V. We show that $v_1 + W, \ldots, v_{n-m} + W$ are linearly independent and span V/W.

Suppose $b_1(v_1 + W) + \cdots + b_{n-m}(v_{n-m} + W) = 0 + W$. Then $(b_1 v_1 + \cdots + b_{n-m} v_{n-m}) + W = 0 + W$. Therefore, $b_1 v_1 + \cdots + b_{n-m} v_{n-m} \in W$. Since $w_1, \ldots, w_m$ is a basis of W, there exists $a_1, \ldots, a_m \in F$ such that $b_1 v_1 + \cdots + b_{n-m} v_{n-m} = a_1 w_1 + \cdots + a_m w_m$. Since $w_1, \ldots, w_m, v_1, \ldots, v_{n-m}$ is a basis of V, we have $b_1 = \cdots = b_{n-m} = a_1 = \cdots = a_m = 0$. This shows that $v_1 + W, \ldots, v_{n-m} + W$ is linearly independent.

Let $x \in V$. Since $w_1, \ldots, w_m, v_1, \ldots, v_{n-m}$ is a basis of V, there exists $a_1, \ldots, a_m, b_1, \ldots, b_{n-m} \in F$ such that $x = a_1 w_1 + \cdots + a_m w_m + b_1 v_1 + \cdots + b_{n-m} v_{n-m}$. Then $x + W = (b_1 v_1 + W + \cdots + b_{n-m} v_{n-m}) + W$. This shows that $v_1 + W, \ldots, v_{n-m} + W$ span V/W. $\square$

Remark 3.17. In view of Corollary 3.8, we see that if U is the complement of W in V then $\dim U = \dim V/W$ and if $u_1, \ldots, u_k$ is a basis of U, then $u_1 + W, \ldots, u_k + W$ is a basis

V/W. We can think of V/W "the same as" U. More precisely V/W is isomorphic to U. We will discuss isomorphic in the next section. The thing is V/W does not depend on a choice of a basis and the construction also works for infinite dimensional vector spaces.

3.3 Linear transformations

The key object we study in linear algebra is linear transformation which is a map between vector spaces preserving the vector space structure.

Definition 3.18. Let V, W be vector spaces over F. A map $T : V \to W$ is a **linear transformation** (**linear map/mapping/operator**) if for any $x, y \in V$ and $c \in F$ we have

$$T(x + y) = T(x) + T(y) \text{ and } T(cx) = cT(x).$$

Remark 3.19. If we take $k = 0$ in the second equation then we have $T(0) = 0$.

Example 3.20 (The identity map)**.** The map $\mathrm{id} : V \to V$ defined by $\mathrm{id}(x) = x$ for any $x \in V$ is a linear transformation called the **identity map**.

Example 3.21 (The zero map)**.** The map $0 : V \to W$, $0(x) = 0 \in W$ for any $x \in V$ is a linearly transformation called the **zero map**.

Example 3.22. The most important linear transformation in this course is $T : F^n \to F^m$, $T(x) = Ax$ where a matrix $A \in M_{m\times n}(F)$ is multiplying a vector $x \in F^n$ is defined by

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$

The matrix multiplication follows the rule "row by column". We will see in the next section that it is a natural consequence of linearity.

Example 3.23. $T : F^n \to F^m$, $T(x_1, \dots, x_n)^t = (x_1, \dots, x_r, 0, \dots, 0)^t$ for some $0 \le r \le \min(m, n)$ is a linear transformation. What the map is doing is first forgetting the variables $x_{r+1}, \dots, x_n$ and then adding $(m - r)$-0 components.

Example 3.24. For $0 \neq b \in F^n$, the map $T : F^n \to F^n$, $T(x) = x + b$ is not a linear transformation since $T(0) = b \neq 0$.

Example 3.25. $T : F \to F$ is linear if and only if $T(x) = ax$ for some $a \in F$. $T(x) = T(x1) = xT(1) = ax$ where we write $T(1) = a$.

$T : \mathbb{R} \to \mathbb{R}$, $T(x) = x^2$ is not linear since $T(1 + 1) = 4 \neq 2 = T(1) + T(1)$.

Example 3.26. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable map. The differential of f at $x \in \mathbb{R}^n$ $df_x : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation. Recall that the differential acting on a vector $v \in \mathbb{R}^n$ is defined by the directional derivative of f in the direction of v

$$df_x(v) = \lim_{h\to 0} \frac{f(x + hv) - f(x)}{h}.$$

By the chain rule, $df_x(v) = Df(x)v$ where the right hand side is the Jacobian matrix

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

multiplying the vector $v \in \mathbb{R}^n$.

Example 3.27.

$$T : C(\mathbb{R}) \to \mathbb{R}, \quad T(f) = \int_a^b f(x)dx$$

is a linear transformation

$P : C^1(\mathbb{R}) \to C(\mathbb{R})$, $P(f)(x) = f'(x)$ is a linear transformation.

$Q : C(\mathbb{R}) \to C(\mathbb{R})$, $Q(f)(x) = xf(x)$ is a linear transformation.

The two linear transformations P, Q are the fundamental operators in quantum mechanics.

Also the operator T and P are the main operators studied in calculus/analysis. Of course, linearity is probably the less important property considered in that subject.

3.4 Operations on linear transformations

Definition 3.28 (Addition and scalar multiplication of linear transformations)**.** Let $T, S : V \to W$ be linear transformations and $c \in F$. The maps $T + S$ and cT are defined as: for any $x \in V$,

$$(T + S)(x) = T(x) + S(x), \quad (cT)(x) = cT(x).$$

It is easy to check that $T + S$ and cT are both linear transformations.

Let $\mathcal{L}(V, W)$ denote the set of all linear transformation. Then with the addition and scalar multiplication above, one can also check $\mathcal{L}(V, W)$ is a vector space. We also write $\mathcal{L}(V) = \mathcal{L}(V, V)$.